Bimal Timilsina

Machine Learning Engineer with a focus on Fine-tuning
and building LLM Applications with RAG

timilsina.bimal04@gmail.com
(977) 986 7232628
github.com/timilsinabimal
bimaltimilsina.com.np

Relevant Experience

ML Engineer · Otermans Institute Sept 2023 — Present

- Building RAG systems with Commercial Large Language models like OpenAI GPT4, Anthropic Claude and Google Gemini using langchain and FastAPI
- Research Open-Source LLMs for the specific use case of Company
- Fine-tune Large Language Models to use with RAG systems
- Reduced the latency of LLM inference on AWS by more than 100%
- Lead the development and architecture of Projects

ML Engineer · Rippey AI Feb — July 2023

- Helped Logistics industry respond to quote and tracking emails in less than a minute using email automation system
- Improved Internal email classification system having three classes by 15%

Associate NLP Engineer · EKbana Solutions Jan 2022 — Jan 2023

Trainee NLP Engineer Jan 2022 — Jan 2023

- Extracted information from Invoices such as invoice no, poid, vendor details, detected and recognized table and cell structure present in the invoice semi automatically with more than 90% accuracy
- Extracted information from insurance documents using OCR methods and applying different parsing methods to gather name, company, insurance number and other information.
- Researched on Nepali Text-to-Speech Model using Tacotron2 and waveglow model and achieved nearly human level results in Nepali Language.
- Researched and Worked on Nepali word variation model using Word embeddings, Cosine Similarity, Edit distance and Jaccard distance.
- Worked on Nepali Keyword Extraction, Speech Recognition, Speech Diarization models.

System Administrator · KU Boys Hostel March 2021 — March 2022

- Helped 250 students with Internet problems
- Managed Switches, Routers, DHCP, SSH and Ubuntu Servers
- Collaborated with ISMS to solve Internet issues

Skills

Programming Language

Python, JavaScript (Basics), HTML, CSS, MySQL, Mongodb

Library and Frameworks

PyTorch, Langchain, Huggingface Transformers, Scikit-learn, FastAPI, NLTK, Numpy, pandas, Matplotlib, Selenium, BeautifulSoup, Django, Spacy

Tools and Platforms

Git, GitHub, Docker, AWS, Jupyter, Conda, Jira

Selected Projects

Background Blurring with Image Segmentation

Implemented and Trained SegNet, UNet, PSPNet and DeepLab models from scratch with 97% Accuracy.

Paraphrase and Semantic Similarity Detection

Trained LSTM, GRU and ConvNet siamese network on SNLI corpus with 76% F1 score.

Education

Kathmandu University 2017 — 2021

Bachelor of Engineering in Computer Engineering with Machine Learning and Natural Language Processing(NLP) Minor

Interests

Coding and Learning, Science Fiction and Thriller Novels and Movies, Traveling and Hiking